DOCUMENT RESUME

ED 133 345 TM 005 931

AUTHOR Fincher, Cameron
TITLE Standardized Tests, Group Differences, and Public Policy.
PUB DATE 20 Oct 76
NOTE 37p.; Paper prepared for an invited lecture in the College of Education, Ohio State University

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Ethnic Groups; Intelligence; Nature Nurture Controversy; *Public Policy; *Racial Differences; *Standardized Tests; Test Bias; *Testing Problems
IDENTIFIERS *Group Differences

ABSTRACT

The paper considers the controversies of standardized tests and group differences as they relate to the formation of public policy. The overlapping or confounding issues of standardized tests and group differences have long since become matters of public policy. Neither separately nor jointly can the issues be resolved within the confines of professional, technical, or administrative judgment. Unless suitable policy alternatives can be identified and deliberately chosen, the controversies that currently prevail in test usage and group measurement will not be resolved. After discussing the nature of the controversy, heredity versus environment, the importance of IQ, and the standard in testing, the author suggests six alternatives, although not exhaustive, which may be discussed as policy solutions that would presumably resolve the issues and controversies involved. These alternatives include: (1) a moratorium on testing, (2) the adjustment of group differences, (3) new rationales for testing, (4) the teaching of intelligence, (5) alternative schooling, and (6) adaptive treatment. Each alternative is discussed in some length. (RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED133345

STANDARDIZED TESTS, GROUP DIFFERENCES, AND PUBLIC POLICY

Cameron Fincher
Institute of Higher Education
University of Georgia

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM005 931

Paper prepared for an invited lecture in the
College of Education, Ohio State University
on October 20, 1976

STANDARDIZED TESTS, GROUP DIFFERENCES, AND PUBLIC POLICY

If policy issues in education were judged by the intensity of feelings they arouse, the uses of standardized tests and the measurement of group differences would rank quite high. Few issues can match the controversy surrounding the measurement of racial-ethnic group differences on standardized tests for emotional intensity and specious arguments.

The controversy continues to dominate an incredible amount of professional thinking, and public debate concerning various facets of the controversy continue to receive coverage in the popular press. The "IQ Myth" has been the subject of a nationwide television documentary and may have been handled in precisely the manner that would ensure its continued controversy. Federal courts and administrative agencies continue to deal with the problems of standardized tests and group differences in personnel selection, promotion, and transfer; state boards of education, university systems, and other educational agencies continue to debate the educational value of standardized tests and their fairness to racial and ethnic groups; and varying forms of anathema continue to be pronounced on the ideological heresies of Jensen, Herrnstein, and Shockley.

The purpose of this paper is to consider the controversies of standardized tests and group differences as they relate to the formation of public policy. Although the two issues have separate and distinct features, their abrasive interaction constitutes one of the major problems of education in a pluralistic society. The persistence

of certain group differences on standardized tests is perceived by many critics as spurious, and it is the continuance of those differences that spur their criticism of tests. Without racial-ethnic differences, the testing controversy would be less intense and might be resolvable. Without the limitations and disadvantages of standardized tests, the problems of racial differences might be considered more easily in terms of their educational implications. Standardized tests have been developed within a context of individual differences; racial and ethnic differences are debated in a context of group interest that now suggests the replacement of a psychology of individual differences by a sociology of groups (Glazer, 1975).

A major premise of the paper is that the overlapping or confounding issues of standardized tests and group differences have long since become matters of public policy. Neither separately nor jointly can the issues be resolved within the confines of professional, technical, or administrative judgement. Unless suitable policy alternatives can be identified and deliberately chosen, the controversies that currently prevail in test usage and group measurement will not be resolved.

For purposes of this paper, policy may be defined as an emerging consensus of opinion that permits or encourages specific lines of public action. Policy represents a pragmatic concern for consequences, outcomes, or results as well as a directive stimulus for public action and can be interpreted glibly as a public answer to the question, "What do we do about it?" To paraphrase the logical empiricists, unless there is a criterion by which to judge the answer, the question may be meaningless. Unless policy alternatives can be defined for the controversies of standardized tests and racial-ethnic group differences, we will continue to debate the issues in a scientific, professional, and ethical vacuum.

The Nature of the Controversy

The problems and issues concerning standardized tests and group differences may be considered in several ways. The issues generating the most heat are those concerning the influence of hereditary or genetic factors in test performance and their consequent determination of measured group differences. This controversy has been much revived by the publications of Arthur Jensen (1969), Richard Herrnstein (1973), and William Shockley (1972). It has been perpetuated by critics who perceive the public school as serving a power structure committed only to its own perpetuation. While these critics may vary in their acceptance or rejection of genetic factors, they are disposed to see in standardized tests a perpetuation of class or cultural bias that contradicts certain basic premises of a democratic society. It is their contention that standardized tests reflect the values, knowledge, and attitudes of a middle class and are not adequately attuned to the motives, viewpoints, and prior education of culturally different groups. Some critics speak from a pronounced ideological base that would preclude the uses and applications of standardized tests in education, industry, and government.

Heredity Versus Environment

The nature/nurture issue in psychology and education has consumed volumes of print. While most participants in the argument would avoid an either-or position, maintain a semblance of detachment and objectivity, and refrain from ideological attacks or ad hominem arguments, the tragedy may be that almost all have been unable to do so. The issue apparently arouses feelings of great intensity and discussants are evidently trapped into taking and trying to hold positions that are

intellectually indefensible.

In an earlier debate, Walter Lippmann (1976) could not refrain from referring to items on the Stanford-Binet as "stunts" and built his argument as much on ridicule as his pertinent and quite justified suspicions that the testing movement had taken some undesirable turns. This provoked from Lewis Terman (1976) an emotionalized defense that missed many criticisms he himself would have offered at another time and place. As a humanist, perhaps, Lippmann was incensed that anyone could be so presumptuous as to attempt the measurement of "intelligence". Terman implied that since Mr. Lippmann knew so much, he had an obligation to abandon his role as critic and enter "this enchanting field of research".

The dangers of entrapment are seen in the Jensen affair. Accurate estimates of the readership of Jensen's original article in the Harvard Educational Review are not available, but it is quite safe to infer the article has had more critics than readers. Those who have read the article may fault Jensen more for his polemical style than for anything he said about racial-ethnic group differences. The fact that he stated his conclusion first makes his argument sound as if it were deduced from first premises. As Cronbach (1975) has pointed out, Jensen intended to speak for open-mindedness and continued research but was trapped into defending positions he had not taken previously.

Much the same reaction can be seen in the Herrnstein incident. The original article in the Atlantic Monthly was accepted only after Herrnstein had agreed to certain changes suggested by the editor. The point may be that in both Jensen's and Herrnstein's articles, there was an editorial desire for a controversial presentation of what should

have been a researchable hypothesis rather than a debatable topic. Noam Chomsky (1976), in one reaction, suggests that Herrnstein's IQ may be fallacious; Herrnstein (1976), in rebuttal, wonders "Whatever happened to Vaudeville?" Given the final word by an editor, Chomsky repeats his original criticism that while Herrnstein's position is both trivial and fallacious, it will surely be used by racists to justify discrimination.

The tendency to accuse the opposition of an ideological bias is particularly strong in the controversy. Leon Kamin (1974) infuses considerable moral indignation in his attack on the ideological biases of the pioneering generation of test constructors and users. His criticism of Sir Cyril Burt's work is unusually devastating and his sensitivities to another generation's methodological weaknesses are quite acute. But Kamin may be more quickly trapped into an either-or position than he recognizes, and his own ideological preferences are more dominant than his own "reasonably prudent man" would expect.

The influence of ideology is seen even more easily in the arguments that while intelligence may have an hereditary basis, it is not a major determinant of success in the economic, social, and political spheres of life. Christopher Jencks and his associates (1972) have estimated that 45 percent of the total variance of IQ scores may be attributed to heredity, 35 percent to environment, and the remaining 20 percent to an interaction of the two. More important, it is their contention that differences in intelligence are not due to factors controlled by the schools and do not account for much of the variances in occupational status and income.

Under the auspicies of the Social Science Research Council, Loehlin, Lindzey, and Spuhler (1975) have sought to define an intellectually

honest position on racial differences in intelligence. They contend that no issue in the social sciences has proven to be more intrusive than the question of assessing the relative importance of the biological and environmental determinents of behavior. They proceed on the premises that: (a) genetic does not mean unchangeable, (b) equal does not mean identical, (c) the unity of general intelligence has not been fully established, (d) there are different meanings of race, (e) there are disagreements about heritability, and (f) scientific knowledge itself has a probabilistic nature.

Their review of the research literature confirms that many studies are poorly designed, executed, and reported and that the political and social preferences of the investigators often bias their interpretation of data. They conclude that:

1. Racial-ethnic differences and socioeconomic differences are sufficiently distinct to demand separate treatment.

2. IQ scores correlate with socioeconomic indices within minority groups as well as in the majority group.

3. Social change can affect average levels of test performance in various sub-populations, but the average group differences do not seem to be particularly responsive to such environmentally induced changes.

4. Stimulating environments can have a substantial effect on the measured IQ's of young children but it is not clear which environmental changes are effective or how permanent the resulting IQ changes are.

5. There is some evidence of differences in patterns or profiles of ability but it is not known to what extent these differences are due to cultural or genetic factors.

6. The majority of variation in either patterns or levels of ability lie within racial-ethnic groups and not between them; neither race nor social class should be used as a predictor of individual's performance on tests of intellectual ability.

Loehlin, Lindzey, and Spuhler emphasize that differences in intelligence reflect: (a) inadequacies and biases in the tests themselves,

(b) differences in environmental conditions, and (c) genetic differences among the groups. These factors are not necessarily independent and may interact. Current evidence permits a wide range of positions to be taken concerning the relative weight of the three factors and, regardless of relative importance, there is no doubt that within group differences exceed in magnitude the average differences between such groups. They are aware that these conclusions are limited but do not believe that scientific evidence at the present time justifies stronger ones.

In considering the social and public policy implication of genetic differences, these authors contend quite strongly that no public policy implication follows from scientific conclusions alone. Public policy is a function of both empirical data and social values. They believe that the following social values should command assent: (a) group membership should not be the occasion for feelings of shame or inferiority, (b) members of all groups should have full opportunity for economic reward, social acceptance, and political participation, and (c) members of all groups should have equal access to social benefits and services.

How Important Is IQ?

The argument that intelligence contributes very little to individual economic success is presented at great length by Bowles and Gintis (1976). They attack what they believe to be a technocratic-meritocratic ideology and contend that the public schools legitimate economic inequality. The nation's school systems presumably foster the belief that economic success is dependent upon technical and cognitive skills which are provided in an efficient and equitable manner based on merit. While

there is a relationship between education and economic success, this relationship cannot be accounted for in terms of cognitive achievements. They believe an educational meritocracy to be largely symbolic.

Loehlin, Lindzey, and Spuhler conclude that while intelligence is of genuine importance in most areas of success, it is far from being all-important in determining what life will be like for most people. They suggest a number of interesting studies on racial-ethnic differences in intelligence but raise the question of urgency. It is their conclusion that for studies of educational methods and policies, research into individual differences is more critical than research into group differences. The urgency of research into possible eugenic or dysgenic trends is a matter of opinion. Their position may be contrasted with that of Block and Dworkin (1976) who conclude after a lengthy survey that while research into racial differences should not be prohibited, individual scientists, because of the political climate at this time, should voluntarily refrain from the investigation of racial differences in intelligence. Loehlin, Lindzey, and Spuhler conclude their discussion with the contention that a humane and enlightened public policy should not be bound by either hereditarian or environmentalistic dogmas. Public policy should be responsive to the fact that individual variation greatly exceeds group differences and that while intelligence is an empirically significant variable, it is not everything. Intelligence may be socially defined in terms that are not identical with performance on IQ tests. Intellectual achievements are not merely a matter of ability but reflect motivational, temperamental and opportunity factors as well. The relationship between intelligence and social rewards is at best only moderate. Finally, they close with the statement that moral and political

questions never have had scientific answers.

The Standard in Testing

The pervasive dissatisfaction with standardized testing stems from misgivings concerning the norms or standards against which individuals are compared (Messick, 1975). The basic premise of psychological and educational measurement either goes unnoticed or becomes highly questionable when brought into public debate. Yet, it is the absence of an absolute standard that prompted the development of standardized tests and their success in providing a comparative base for decisions and judgements concerning human behavior. It is well to recall that "standard scores" provide a means of comparing individuals with a specific norm group and were not designed for the comparisons of groups as such. It is the individual's test performance that is adjusted for the group's central tendency and variation.

The use of standardized tests for group comparisons is the core of the controversy. There is the criticism of inherent bias favoring the groups upon which the tests were originally developed and "normed"; there is the contention that the norms do not reflect what teachers teach and students learn; there is the apprehension that standardized tests will be used for judging teaching effectiveness; there is dismay at the extent to which testing is a part of the educational scene; and there is a general distrust of the value of standardized tests in making educational decisions and judgements.

The controversy surrounding standardized tests is further exacerbated by test score decline over the past fifteen years (Munday, 1976; Harnischfeger & Wiley, 1976; Flanagan, 1976); their use as criteria in sociological and econometric analysis of schooling (Coleman, et al,

1966; Averch, et al, 1972; Levine, 1976); the difficulties of debating their technical features in courts of law (Lennon, 1966); and their suspected abuse in classifying and sectioning students for learning and teaching purposes (Findley & Bryan, 1971; Mercer, 1974).

Policy Alternatives

If policy is accepted as the criterion by which to judge group differences on standardized tests, several alternatives may be seen in the recommendations and suggestions of various critics debating the testing problems and issues of the past several years. Some testing specialists would regard the cures and remedies of testing differences as homeopathic and insist on further study and research of the problem, holding social consequences and costs in abeyance until the issues can be resolved in a purely scientific manner. Other critics would insist on more instant solutions that would alter or modify radically the nature of tests, schools, and society in general. Although not exhaustive, six alternatives may be discussed as policy solutions that would presumably resolve the issues and controversies involved. These alternatives include: (a) a moratorium on testing, (b) the adjustment of group differences, (c) new rationales for testing, (d) the teaching of intelligence, (e) alternative schooling, and (f) adaptive treatment.

A Moratorium on Testing

Some critics of testing would resolve the problems and issues of group differences by an immediate abandonment of testing. Either through professional restraint, legislative action, or court rulings; they would insist that tests not be used for purposes of selection, placement, and evaluation until tests meet standards of validity and reliability that are believed to be essential. The contention is that

standardized tests are culturally unfair and improperly classify members of minority groups. More important is the insistence that standardized tests are inherently biased in the assumptions on which they are based. Distinctions between intelligence or aptitude tests and achievement tests are spurious. Both kinds of tests serve to maintain the differential status of racial-ethnic groups in a society that is committed to the continued dominance of a single cultural tradition (Mercer, 1974; Williams, 1970; Houts, 1976).

Efforts to counter the call for a moratorium have usually taken the tact that a distinction must be made between a charge of inadequacy on the part of all tests and the abuses of testing practices. A more telling argument is presented in terms of the replacement possibilities. The alternative to testing is seen as forms of subjective appraisal that have been tried in the past and fully established as wanting. The advantages of testing as a means of systematic inquiry, objective data, and useful information would be the baby discharged with the bath (Messick and Anderson, 1970).

There is further reason, however, to believe that a complete abandonment of testing is not a viable policy alternative. It is well to recall that the intense criticism of testing in the early Sixties resulted not in a reduction of the amount of testing in industry and education, but came at a time when other policy issues and decisions were dictating an expanded use and application of standardized tests. (Holmen and Docter, 1972; Brim, Glass, Neulinger, and Firestone, 1969).

Cleary, Humphreys, Kendrick, and Wesman (1975) have pointed out that the fair use of standardized tests does not depend on a solution to the heredity-environment problem. Tests have been developed under ground rules that make no casual assumptions as to test performance. All

tests measure current performance and not innate capacity.

These authors suggest that one alternative may be identified as random selection or selection by lottery. Such a procedure was used in the admission of the first class to Federal City College in Washington, D.C. and implies that a lottery may be as "fair" as other selection methods. Other alternatives are: (a) the use of prior experience as determined in work histories or biographical inventories, (b) the use of demographic categories as used in quotas and schedules, and (c) the many forms of subjective evaluation, such as interviewing, that have been used in the past. For educational purposes, the best alternative to testing remains the student's previous academic achievement as represented by grades. But Cleary, Humphreys, Kendrick, and Wesman contend that the abandonment of tests would mean the abandonment of those functions supported by tests and that we would be left without objective appraisals. They also point out that any alternative to testing would have to meet the same criterion of fairness that tests themselves are criticized for not having.

The Adjustment of Group Differences

Many professionals and testing specialists apparently believe that the social or cultural bias in standardized tests can be controlled through statistical or methodological adjustments. These advocates would take the limitations of tests as they now exist and adjust, through regression analysis, the group differences that provide a source of disagreement. Such corrections would be made in keeping with concepts of fairness or compensatory justice for previous discrimination against ethnic or minority groups (Fincher, 1973).

A basic premise of such efforts is the notion that group differ-

ences on standardized tests are a technical problem to be resolved from within the profession itself. Group membership would be treated as a moderator variable that should actually sharpen the prediction made of future performance by two or more groups (Bartlett & O'Leary, 1969). Although not treated as a moderator variable as such, sex has long been used in testing as a basis for separate norms, prediction equations, and selective treatment. An effort is often made to control sex bias in the construction of test items by not including items which favor one or the other. Sex differences are openly acknowledged in standardization procedures and then controlled by the development of two sets of norms, presented in most cases with norms for the combined groups. Biases in the selection process itself are further controlled by the use of separate equations that would openly consider the sex of the individual applicant.

The correction of social or cultural bias through differential validation has not been highly successful in either employment or education. The technique is apparently limited because of the small number of subjects available in industrial validation studies and because of the frequent difficulties in identifying applicants by race or ethnic group. Such studies were often needed at a time when other rules and regulations prohibited the identification of applicants by race or ethnic grouping.

Although differential validity is explicitly regarded as a policy alternative in federal agency guidelines prepared under Title VII of the Civil Rights Act, its relevance for public policy has not been consistently perceived by courts interpreting legislative intent, and there is an increasing suspicion that its use in predictive studies may constitute a form of reverse discrimination which courts are

beginning to find nonpermissible under the Fourteenth Amendment.

Other efforts to define test bias, or its counterpart, test fairness, also have been unsuccessful. Although a variety of ingenuous methods have been suggested, the inherent weakness of such methods have eventually been brought forth by specialists seeking a better way of defining, and thereby controlling, test bias. Examples may be seen in Cleary's (1968) attempt to define bias in terms of the accuracy of prediction that is possible, Thorndike's (1971) effort to ensure comparable group outcomes, Einhorn and Bass' (1971) attempts to establish equal probability of success, and Cole's (1973) appeal for a method that would equalize the probability of selection.

Darlington (1971) represents a broader view that would take policy preferences into explicit consideration and openly acknowledge that if test bias is to be eliminated, test validation must not only be empirical but test usage must be rational. Peterson and Novick (1974) would go further and insist upon an equity of expected utilities that are derived within the framework of a general decision-theoretic formulation. In doing so, public policy would dictate what is to be accomplished by the decisions made.

Other observers have questioned the success of both differential validity and methodology definitions of test bias (Fincher, 1975a; Schmidt and Hunter, 1974). There is the contention that both traditional test theory and current testing rationales are contrary to the adjustments that are attempted. Lord (1967) is emphatic in the possibility that group adjustments may not be permissible in an ex post facto manner. Cronbach (1976) suggests that the problem is indeed one in which psychometrics and political philosophy meet.

New Rationales for Testing

Other critics, working both within and without the field of testing, have seen the abuses of standardized tests as stemming from the basic premises of testing. They believe the abuses can be eliminated through new departures in the philosophy, purposes, and rationales of testing and advocate a systematic concern with different forms and procedures. These critics would modify the uses and applications of testing by shifting the predominant emphasis from selection and predictive validity to the facilitation of learning and teaching, as well as new practices and implications.

The most prominent of these critics is Ralph Tyler (1974). It is his contention that tests have been developed for particular purposes with assumptions and techniques appropriate to the times. With the change of times, tests must be used for new purposes and must, therefore, operate on different assumptions with techniques that are suitable. Testing in the seventies must be relevant to the extension of minority group rights and should serve our needs to evaluate programs and instructional methods, as opposed to individual capacities and differences. Tyler argues most persuasively that standardized tests were developed in an era when a psychology of individual differences was dominant and served primarily to rank-order students. Test scores were used for classifying, placing, and sectioning students for traditional modes of instruction. The needs of the postsixties call for instruments that measure educational outcomes, as opposed to initial aptitudes and abilities, and enable us to monitor, evaluate, and improve what students learn and what teachers teach.

Support for Tyler's position may be seen in the emergence of evaluation research as a professional interest and specialty (Phi Delta

Kappa National Study Committee on Evaluation, 1971; Tyler, 1969; Anderson, Ball, Murphy, et al, 1976; Suchman, 1967), the strong push for criterion-referenced tests, as opposed to the norm-referenced features of traditional tests (Glaser and Nitko, 1971; Knapp, 1974), the pervasive concern with statewide assessment at the public school level (Ebel, 1974; Fincher, 1975b), and the emphasis placed on National Assessment of Educational Progress (Johnson, 1975; Mellon, 1975).

The policy implications of these trends and developments are seen most clearly perhaps in the uses and applications of criterion-referenced tests. The development and use of such instruments signify a disenchantment with obstinate group differences that are a continuing source of embarrassment in a society committed to educational values. An explicit assumption of the movement is the belief that relative standards constitute a questionable comparison of individuals and groups with the sole benefit being an ordering or ranking of students. Normative approaches indicate where they stand with reference to some group which is often vaguely identified or inappropriate. The appeal is for absolute standards that are properly specified in educational or learning terms and which the individual can reach without regard to measures of group performance.

Robert Ebel (1975) is among critics who believe criterion-referenced tests to possess the appeal of novelty and innovation but not freedom from technical problems or policy difficulties. Not only is the problem one of specifying and sampling domains of knowledge but there is the difficulty of setting standards or levels of performance which will be rational and acceptable. Ebel contends the assumptions of criterion-referenced testing are mistaken. Since both criterion-referenced and norm-referenced tests must be designed to differentiate

levels of achievement, the need to make comparisons among individuals cannot be avoided.

It is significant, therefore, that results of the National Assessment of Educational Progress give much the same picture as that revealed over a longer period of time by standardized tests. Although criterion-referenced in terms of individual performance and reported as the percent passing or succeeding on specific tasks, the results of National Assessment display the same regional, race, and sex differences seen in norm-referenced measures. Also relevant is the finding by Harnischfeger and Wiley (1976) that test score decline is evident not only in scholastic aptitude tests but also in educational achievement tests. Although this finding may support the argument that there are no differences between intelligence or aptitude tests and achievement tests, it is nonetheless indicative of a more general problem than the one posed by the advocates of criterion-referenced tests.

Yet, the policy implications of the new rationales for testing remain much the same. Current testing practices reflect a rationale that has historical roots and traditions that may not be adequate for present problems and issues in education (Anderson, 1972; Samuda, 1975). As Levine (1976) has pointed out, standardized achievement tests had profound significance in an era when both psychology and education aspired to scientific status and acceptance. Just as the testing movements of that day reflected the aspirations of its participants, so they reflected the social values and norms current at that time. There is little doubt that standardized tests, with or without a "modern rationale", reflect the values and policy preferences of dominant interest groups.

The Teaching of Intelligence

One alternative to the nagging persistence of group differences has been suggested as the provision of special educational programs that would eradicate such differences by developing the skills and competencies disadvantaged groups obviously need in taking such tests. The teachability of intellectual and cognitive abilities is a topic much debated in itself. Coupled with the difficulties of unwanted group differences, the topic reveals a great deal about the optimism or pessimism of its opposing discussants. Those who are optimistic about the issue contend that the IQ is clearly related to the perceptual habits, cognitive skills, previous learning, and personal values of the individual rather than to his or her racial background or group membership. Assuming that these characteristics are acquired through experiental learning, it presumably would follow that intelligence itself should be amenable to instruction if that instruction were properly arranged and provided.

Arthur Whimby (1975) may be cited as one advocate who takes an optimistic view of instructional efforts to ameliorate group differences in measured intelligence. He adopts the view that intelligence, whatever it is, is a skill that can, like other skills, be taught and learned. Intelligence consists of problem-solving skills that are generalized from situation to situation. Intelligence tests consist of abstract concepts,general principals, and problem-solving tasks that require abilities or skills presumed to be present, to some degree, in all persons taking the tests. Since intelligence involves the rudiments of learning, it would be an unwarranted concession to conclude that the schools cannot teach what students have obviously learned.

Although Whimby's arguments warrant consideration, they are

countered by a far more pessimistic view of programmatic efforts to deal with educational disadvantages. The criticisms of Headstart, Upward Bound, and other ambitious federal programs of the Sixties do not lend great support to optimistic expectations that group differences are easily overcome. To the contrary, it is the perceived failure of these programs that launched the Jensen affair, created the notoriety of the Coleman report (1966), and spurred those who would release the schools from any further promise of eliminating the social problems of inequality in cognitive abilities and educational outcomes.

At the postsecondary level, the problems of teachability have been aroused by the pessimistic findings surrounding compensatory, remedial, or developmental studies. While there are nuances of differences among the three concepts, all three are an effort to deal with educational disadvantages that have persisted through twelve years of public schooling. The need for such programs is a function of the social movements of the sixties but stems from a longer, more extended effort to make education beyond the high school accessible to larger numbers. Where groups of students have been referred to in the past as marginal or academically handicapped, the federal programs of the sixties produced a pervasive concern for students who were culturally deprived, economically dispossessed, or educationally disadvantaged. Special programs were organized to accomodate large groups of students who had not previously entered college and who were not previously expected to seek entrance to professional, scientific, or technical occupations.

Mulka and Sheerin (1974) have surveyed published studies dealing with programs for the educationally disadvantaged at the postsecondary level. Their findings suggest that most efforts to provide relatively short term compensatory or remedial education have resulted in ambiguous

or discouraging outcomes. Efforts at the postsecondary level have suffered greatly from a lack of clear-cut purposes or a general understanding of what the programs were to accomplish. Many colleges succeeded only in discovering that the problems of the disadvantaged were pervasive, deeply entrenched, and unresponsive to hasty, superficial efforts at amelioration. Too frequently remedial or compensatory courses were treated as a general re-cycling of what the disadvantaged students had not learned in reading, math, and grammar as those subjects were taught at the elementary and secondary levels. The lack of trained diagnosticians and skilled instructors could not be off-set by the use of technological devices and programmed instruction. In general, such courses were encapsulated as either ante-room operations that neither facilitated the student's mastery of traditional subject matter courses nor eradicated the deficiencies in fundamental skills with which he came.

Mulka and Sheerin believe the problems of compensatory education to persist because there is an absence of nation-wide standards for basic competencies in oral and written communication, reading, and mathematics. A major recommendation of their study is an intensified effort on the part of the U.S. Office of Education to establish competency levels for grades one through twelve. Given the difficulties of establishing such standards, it would follow that postsecondary institutions are sorely handicapped in their efforts to correct or alleviate basic deficiencies in short-term programs.

The inevitable criticism of efforts to teach intellectual or cognitive skills is that such "training" is readily seduced into "teaching the test". There remains a widespread feeling that attempts to teach the problem-solving or critical thinking skills are, in some way, "unfair".

This feeling, however, may not be as pervasive as the suspicion that such skills are not actually teachable within the framework of public schooling because of other demands on school resources.

Alternative Schooling

Many critics and observers have wearied of the controversy surrounding inequality and group differences, openly confess a disdain for its continued discussion, and propose radically new paths to adulthood and employment. Some of their premises are neo-romantic in their assumptions about childhood and youth, ascerbic in their rejection of public schooling as a liberalizing force, and scornful of the school's authoritarian stand against creativeness on the part of students. The argument is that schools have failed and must be replaced by other alternatives and options.

The attack on the schools has been articulated by a host of highly vocal critics. George Dennison (1969), James Herndon (1968), John Holt (1970), Jonathan Kozol (1968), and Ivan Illich (1971) have spelled out in pungent terms the dissatisfactions with public schooling as a means of achieving social objectives. The economic benefits of schooling have been severely questioned by Jencks (1972), Bowles and Gintis (1976), and by numerous others who have cast suspicion upon the credentialing and certifying role in which colleges and universities have permitted themselves to slide (Juster, 1975; Solmon and Taubman, 1973). The Panel on Youth (1974) has advocated new approaches to the responsibilities of adulthood and emphasizes, in particular, a systematic alteration of school and work that would change appreciably the traditional roles of the school as preparatory experience for work. Other alternative routes suggested by this panel include organizations of youth

and adults that would eliminate the age segregation that youth are subjected to by our present systems of education.

Supportive positions for the Panel on Youth may be seen in reports by the National Commission on the Reform of Secondary Education (1973), and by the Special Task Force to the Secretary of HEW (1973), as well as the more recent report from the National Manpower Institute, written by Willard Wirtz (1975). Both reports are emphatic in their discussion of new routes, paths, bridges, etc. that would permit youth to gain a better appreciation of work in an industrial society. The intent of these reports is to make both employment and education more satisfying and meaningful, but most assuredly, to bridge the gap between the two. Both would cope with the policy issues of education by specific lines of public action at the federal, state, and local levels.

At the postsecondary level an intense interest has been shown in alternative learning or nontraditional study. The concept of nontraditional study has been endorsed by the Carnegie Commission on Higher Education (1973), the Commission on Nontraditional Study (1973), the Panel on Alternate Approaches to Graduate Education (1973), and other groups dealing with the changing demands being made upon institutions of higher education (Committee on Higher Adult Education, 1972; Study on Continuing Education and the Future, 1972). Each of these groups expresses a concern with alternative systems of learning and other kinds of informal learning that does not fit into the customary mold of formal instruction. Each group is reacting to the excessive rigidity that is seen in traditional systems of higher education, the definite possibility of a declining enrollment in the college-age population, and other changes in societal values.

While national panels and task forces concerned with alternative

paths have sought a constructive, somewhat incremental approach to public policy, the conclusions and recommendations of Bowles and Gintis (1976) are indicative of the radical approach other critics of education would take. Bowles and Gintis conclude that neither intelligence nor educational achievement are important determinants of success in the economic sphere. They are extremely dubious that formal education can reduce the inequities they perceive in family income, occupational status, and other material indices of success. Their alternative to the failures of the school is a radical redesign of society. School reform is essentially meaningless in a society committed to corporate power and a capitalistic economy. It is their contention that our schools have never been "above politics" but openly serve to perpetuate the manpower needs of corporate industry.

The sifting and sorting of the various proposals for school reform will occupy the time of educational thinkers for several years to come. While there is doubt that the majority of such recommendations have strongly dented the consciousness of the general public, they do represent vigorous debate within the arena of public policy. General dissatisfactions with the public schools find a specific focus in the more tenable issues of racial-ethnic group differences, test score declines, grade inflation, and increasing costs. Public alarm is easily aroused with the publicity given the job market for college graduates who once would have found jobs waiting in managerial and technical occupations. The inflationary costs of education, the militancy of teachers, and the obvious organizational complexities of public institutions at both the secondary levels are rallying points for those who fear a waning commitment to education and the values it is presumed to preserve.

As a policy alternative to group differences on standardized tests,

the advocacy of alternative schooling is a capitulation to the difficulties of resolving the issues within the present framework of public education. The policy consequences of some actions, if implemented, would signify that only other systems and institutions can successfully resolve these issues.

Adaptive Treatment

Contrary to those who would seek alternative schooling as a policy resolution to educational issues, there are others who would actively alter the treatment that students presently receive in the public school setting. Policy alternatives in this effort would not be based on a disenchantment with public schooling but upon a critical assessment of internal weaknesses and limitations. Methods of instruction would be modified to permit and encourage a more personalized or individualized form of teaching; instructional materials, curricula, and academic programs would be altered to meet individual needs and interests of students in a more systematic manner. A basic premise of this approach is clearly that individual differences remain an important instructional challenge. There is an awareness of cultural pluralism and diversity as well as continuing faith in those technological approaches and procedures that have proven successful in the past. The gist of this approach would be to adapt the instruction of individuals to their instructional needs and their individual capacities and abilities.

The advocacy of aptitude/treatment interaction, person/environment interaction, or trait/treatment interaction is based on a statistical sophistication of how variables in a learning-teaching situation are apt to be dependent upon each other in peculiar ways (Berliner and

Cahen, 1973; Salomon, 1972; Walsh, 1973). There is a suspicion that education does submit to formalization as a hypothetic-deductive system but requires a sustained searching-out of working hypotheses that permit students to learn more effectively in situations where teachers, subject matter content, environmental conditions, and other situational variables are interacting variables.

Some advocates would encourage a dramatic shift from models of prediction and selection to models of placement and exemption (Willingham, 1974). Such a shift would remove standardized testing from the strait-jacket in which it has been cast and permit a more constructive use of tests in the educational progression of students. Individual differences would be accommodated through a variety of administrative and instructional devices designed to assist both the individual choices of students and institutional or program decisions that must be made by administrators and faculty. More important, perhaps, the approach would require that educational programs be considered within the large framework of the system of which they are a part.

Adaptive treatment would include: (1) assignment, (2) selection, (3) placement, and (4) exemption, depending upon characteristics of the students, the educational programs, and the institutional setting. Placement and exemption models, in particular, would place students in programs adapted for their individual differences. Instructional systems would be developed for personalized purposes and permit students to proceed at a pace determined by their own objectives, needs, and motives (Cross, 1976).

The advocacy of adaptive treatment often suggests a technological fix for a complex social and cultural problem. Yet, the approach derives primarily from premises implicit in the tribute the American society

has always paid education. It is the compelling notion that we should teach students according to their interests, needs, and abilities -- an assumption often made by public education. It is relevant that the notion of adaptive treatment is at the core of Jensen's criticism of compensatory education. His argument could have been interpreted as the conclusion that the ways in which we have been teaching minority group students have not been successful. His corollary is clearly that we need to try different modes of instructions, ones from which the students can more obviously benefit (Jensen, 1973).

How successful adaptive treatment can be as educational policy remains to be seen. The interactions between individual differences and instructional methods may be more complex than first suspected, and the larger effort may be seen in terms of its experimental demands rather than its educational implications. Hunt (1975) has suggested that the paradigm of aptitude-treatment-interaction has not been adequately tried and believes it might be reinstated in terms of Lewin's behavior-person-environment model. Tobias (1976) suggests that ATI's may be highly specific, may vary for different content areas, and may result in limited theoretical and practical utility. Yet, if the paradigm includes prior level of achievement, it suggests an interaction directly related to group differences: The lower the level of prior achievement, the higher the instructional support required to accomplish instructional objectives. Also relevant as interacting variables are the belief systems of teachers and instructors who must work with students from culturally diverse backgrounds (Kirk and Goon, 1975; Rosenthal and Jacobson, 1968).

Cronbach (1975), who initiated the active search for aptitude/treatment interaction, has written that we need to reflect on what it

means to establish working generalizations in a world where most effects are interactive. Empirical relationships change and generalizations decay. Instead of seeking generalizations that we can regard as lawful relations, we should seek working hypotheses that are properly tuned to local conditions, contemporary relationships, and historical circumstances. Our aspiration should be to assess local events accurately and improve short-run control.

A major disadvantage of adaptive treatment, therefore, may be its suitability for experimental inquiry as opposed to formal instruction. Given the climate of group conflict that continues in public education, educators may not be permitted to carry out the forms of inquiry that are needed to establish adaptive methods of instruction. Assignment and selection have not been acceptable modes of dealing with group differences in many past situations. To be acceptable to courts, federal agencies, and adversary groups, adaptive treatment must be perceived as a better method of education and not as educational research.

Implications and Conclusions

An attempt has been made to look at the policy implications of group differences on standardized tests in education. Although the testing controversy has many dimensions and would undoubtedly be an emotionalized issue, there is reason to believe that the persistence of unwanted group differences is at the source of many disagreements concerning the uses and applications of standardized tests. National policy over the past two decades has been directed to social objectives that would ensure equality of access and opportunity to racial and ethnic minorities in the American culture. Strenuous programmatic efforts have been made to provide minority groups with educational

opportunities that would permit their full participation and involvement in national life. Yet, systematic efforts to evaluate the results of those programs have consistently shown group differences that were not anticipated by many proponents of minority group rights.

The efforts to explain those differences have resulted in a needless revival of the heredity-versus-environment issue. Such arguments have distracted attention from what should have been the objective of social policy -- the provision of educational opportunity for individuals of differing capacities, interests, and needs. A psychology of individual differences that once suggested a rational and empirical basis for education has been superseded by a sociology of groups that defines national and public issues in terms of group interest.

Six alternatives have been discussed, but only one would appear adequate for the conflicting needs of a pluralistic and diverse society. Social and educational policy must view education as an adaptive institutional response to changing needs and demands. The continuing controversies over standardized tests and group differences reflect those changing needs and demands but do not facilitate the adaptive response that is needed.

REFERENCES

Anderson, R.C. How to construct achievement tests to assess comprehension. Review of Educational Research, 1972, 42, 145-170.

Anderson, S.B., Ball, S., Murphy, R.T., et al. Encyclopedia of educational evaluation. San Francisco: Jossey-Bass, 1976.

Averch, H.A., Carroll, S.J., Donaldson, T.S., Kiesling, H., & Pincus, J. How effective is schooling? A critical review and synthesis of research findings. Santa Monica: Rand Corporation, 1972.

Bartlett, C.J., & O'Leary, B.S. A differential prediction model to moderate the effects of heterogeneous groups in personnel selection and classification. Personnel Psychology, 1969, 22, 265-268.

Berliner, D.C., & Cahen, L.S. Trait-treatment interaction and learning. In F.N. Kerlinger (Ed.), Review of research in education (Vol. 1). Itasca, Ill.: Peacock, 1973.

Block, N.J., & Dworkin, G. (Eds.). The IQ controversy: Critical readings. New York: Pantheon Books, 1976.

Bowles, S., & Gintis, H. Schooling in capitalist America: Educational reform and the contradictions of economic life. New York: Basic Books, 1976.

Brim, O.G., Glass, D.C., Neulinger, J., Firestone, I.J., & Lerner, S.C. American beliefs and attitudes about intelligence. New York: Russell Sage Foundation, 1969.

Carnegie Commission on Higher Education. Toward a learning society: Alternative channels to life, work, and service. New York: McGraw-Hill, 1973.

Chomsky, N. The fallacy of Richard Herrnstein's IQ. In N.J. Block & G. Dworkin (Eds.), The IQ controversy: Critical readings. New York: Pantheon Books, 1976.

Cleary, T.A. Test bias: Prediction of grades of Negro and white students in integrated colleges. Journal of Educational Measurement, 1968, 5, 115-124.

Cleary, T.A., Humphreys, L.G., Kendrick, S.A., & Wesman, A. Educational uses of tests with disadvantaged students. American Psychologist, 1975, 30, 15-41.

Cole, N.A. Bias in selection. Journal of Educational Measurement, 1973, 10, 237-255.

Coleman, J.S., et al. Equality of educational opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.

Commission on Non-Traditional Study. Diversity by design. San Francisco: Jossey-Bass, 1973.

Committee on Higher Adult Education. Special report on higher education and the adult student. Washington, D.C.: American Council on Education, 1972.

Cronbach, L.J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.

Cronbach, L.J. Equity in selection - where psychometrics and political philosophy meet. Journal of Educational Measurement, 1976, 13, 31-41.

Cronbach, L.J. Five decades of public controversy over mental testing. American Psychologist, 1975, 30, 1-14.

Cronbach, L.J. The two disciplines of scientific psychology. American Psychologist, 1957, 12, 671-684.

Cross, K.P. Accent on learning. San Francisco: Jossey-Bass, 1976.

Darlington, R.B. Another look at 'culture fairness'. Journal of Educational Measurement, 1971, 8, 71-82.

Dennison, G. The lives of children. New York: Random House, 1969.

Ebel, R.L. Educational tests: Valid? Biased? Useful? Phi Delta Kappan, 1975, 56, 83-89.

Ebel, R.L. State testing programs: Status, problems, and prospects. (TM Rep. 40). Princeton: ERIC Clearinghouse on Tests, Measurement, and Evaluation, 1974.

Einhorn, H.J. & Bass, A.R. Methodological considerations relevant to discrimination in employment testing. Psychological Bulletin, 1971, 75, 261-269.

Fincher, C. Differential validity and test bias. Personnel Psychologist, 1975a, 28, 481-500.

Fincher, C. Personnel testing and public policy. American Psychologist, 1973, 28, 489-497.

Fincher, C. The access-placement-retention-graduation of minority students in higher education. (A background paper prepared for a conference on Equality of Access in Postsecondary Education sponsored by the Southern Education Foundation and the Ford Foundation in Atlanta, Georgia), 1975.

Findley, W.G., & Bryan, M.M. Ability grouping: 1970. Athens, Georgia: Center for Educational Improvement, University of Georgia, 1971.

Flanagan, J.C. Changes in school levels of achievement: Project talent ten and fifteen years retests. Educational Researcher, 1976, 5(8), 9-12.

Friedenberg, E.Z. Coming of age in America. New York: Random House, 1963.

Glaser, R., & Nitko, A.J. Measurement in learning and instruction. In R.L. Thorndike (Ed.). Educational measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

Glazer, N. Affirmative discrimination: Ethnic inequality and public policy. New York: Basic Books, 1975.

Harnischfeger, A., & Wiley, D.E. Achievement test scores drop. So what? Educational Researcher, 1976, 5(3), 7-12.

Herndon, J. The way its spozed to be. New York: Simon and Schuster, 1968.

Herrnstein, R.J. I.Q. in the meritocracy. Boston: Little, Brown, 1973.

Herrnstein, R.J. Whatever happened to vaudeville? A reply to professor Chomsky. In N.J. Block & G. Dworkin (Eds.), The IQ controversy: Critical readings. New York: Pantheon Books, 1976.

Holmen, M.G., & Docter, R. Educational and psychological testing: A study of the industry and its practices. New York: Russell Sage Foundation, 1972.

Holt, J. What do I do on Monday? New York: E.P. Dutton, 1970.

Houts, P.L. Behind the call for test reform and abolition of the IQ. Phi Delta Kappan, 1976, 57, 669-673.

Hunt, D.E. Person-environment interaction: A challenge found wanting before it was tried. Review of Educational Research, 1975, 45, 209-230.

Illich, I. Deschooling society. New York: Harper and Row, 1971.

Jencks, C., et al. Inequality. A reassessment of the effect of family and schooling in America. New York: Harper and Row, 1972.

Jensen, A.R. Educability and group differences. New York: Harper & Row, 1973.

Jensen, A.R. How much can we boost I.Q. and scholastic achievement? Harvard Educational Review, 1969, 39, 1-123.

Johnson, S.S. Update on education. Denver: National Assessment of Educational Progress, 1975.

Juster, F.T. (Ed.). Education, income, and human behavior. New York: McGraw-Hill, 1975.

Kamin, L.J. The science and politics of I.Q. New York: Erlbaum Associates, 1974.

Kirk, D.H. & Gibson, S. Desegregation and the cultural deficit model: An examination of the literature. Review of Educational Research, 1975, 45, 599-611.

Kirst, W. & Walker, F. An analysis of curriculum policy-making. Review of Educational Research, 1971, 41, 479-509.

Knapp, J. A collection of criterion-referenced tests. (TM Rep. 31). Princeton: ERIC Clearinghouse on Tests, Measurement, and Evaluation, 1974.

Kozol, J. Death at an early age. Boston: Houghton Mifflin, 1968.

Lennon, R.T. Testimony ... as expert witness on psychological testing. New York: Harcourt, Brace, & World, 1966.

Levine, M. The academic achievement test: Its historical context and social functions. American Psychologist, 1976, 31, 228-238.

Lindquist, E.F. Educational measurement. Washington, D.C.: American Council on Education, 1951.

Linn, R.L. Fair test use in selection. Review of Educational Research, 1973, 43, 139-161.

Lippmann, W. The mental age of Americans. In N.J. Block & G. Dworkin (Eds.), The IQ controversy: Critical readings. New York: Pantheon Books, 1976.

Loehlin, J.C., Gardner, L., & Spuhler, J.N. Race differences in intelligence. San Francisco: W.H. Freeman, 1975.

Lord, F.M. A paradox in the interpretation of group comparisons. Psychological Bulletin, 1967, 68, 304-305.

Mellon, J.C. National assessment and the teaching of English. Urbana, Ill.: National Council of Teachers of English, 1975.

Mercer, J.R. Latent functions of intelligence testing in the public schools. In L.P. Miller (Ed.), The testing of black students. A symposium. Englewood Cliffs: Prentice-Hall, 1974.

Messick, S. The standard problem. Meaning and values in measurement and evaluation. American Psychologist, 1975, 30, 955-966.

Messick, S. & Anderson, S. Educational testing, individual development, and social responsibility. Counseling Psychologist, 1970, 2, 80-88.

Miller, L.P. (Ed.). The testing of black students. A symposium. Englewood Cliffs: Prentice-Hall, 1974.

Mulka, M.J. & Sheerin, E.J. An evaluation of policy related research in postsecondary education for the disadvantaged (Vol. 2 - Technical Report). Washington, D.C.: National Science Foundation, 1975.

Munday, L.A. Declining admissions test scores. (Rep. 71). Iowa City, Iowa: The American College Testing Program, 1976.

National Commission on the Reform of Secondary Education. The reform of secondary education -- a report to the public and the profession. New York: McGraw-Hill, 1973.

Novick, M.R. & Petersen, N.S. Towards equalizing educational and employment opportunity. Journal of Educational Measurement, 1976, 13, 77-88.

Panel on Alternate Approaches to Graduate Education. Scholarship for society. Princeton: Educational Testing Service, 1973.

Panel on Youth of the President's Science Advisory Committee. Youth: Transition to adulthood. Chicago: University of Chicago Press, 1974.

Petersen, N.S. & Novick, M.R. An evaluation of some models for culture - fair selection. Journal of Educational Measurement, 1976, 13, 3-29.

Phi Delta Kappa National Study Committee on Evaluation. Educational evaluation: Decision making. Itasca, Ill.: F.E. Peacock, 1971.

Rosenthal, R. & Jacobson, L. Pygmalion in the classroom. New York: Holt, Rinehart, and Winston, 1968.

Salomon, G. Heuristic models for the generation of aptitude-treatment interaction hypotheses. Review of Educational Research, 1972, 42, 327-343.

Samuda, R.J. Psychological testing of American minorities: Issues and consequences. New York: Dodd, Mead, 1975.

Schmidt, F.L. & Hunter, J.E. Racial and ethnic bias in psychological tests: Divergent implications of two definitions of test bias. American Psychologist, 1974, 29, 1-8.

Shockley, W. Dysgencis, geneticity, raceology: A challenge to the intellectual responsibilities of educators. Phi Delta Kappan, 1972, 53, 297-307.

Shockley, W. Models, mathematics, and the moral obligation to diagnose the origin of Negro I.Q. deficits. Review of Educational Research, 1971, 41, 369-377.

Silberman, C.E. Crisis in the classroom. New York: Random House, 1970.

Society for the Psychological Study of Social Issues. Statement by SPSSI on current IQ controversy: Heredity versus environment. American Psychologist, 1969, 24, 1039-40.

Solmon, L.C. & Taubman, P.J. (Eds.). Does college matter? New York: Academic Press, 1973.

Special Task Force to the Secretary of Health, Education, and Welfare. Work in America. Cambridge, Mass.: MIT Press, 1973.

Study on Continuing Education and the Future. The learning society. Notre Dame: University of Notre Dame, 1972.

Suchman, E.A. Evaluative research: Principles and practice in public service and social action programs. New York: Russell Sage Foundation, 1967.

Terman, L.M. The great conspiracy. In N.J. Block & G. Dworkin (Eds.), The IQ controversy: Critical readings. New York: Pantheon Books, 1976.

Thorndike, R.L. Concepts of culture-fairness. Journal of Educational Measurement, 1971, 8, 63-70.

Thorndike, R.L. (Ed.). Educational measurement (2nd ed.). Washington, D.C.: American Council on Education, 1971.

Tobias, S. Achievement treatment interactions. Review of Educational Research, 1976, 46, 61-74.

Tyler, R.W. (Ed.). Educational evaluation: New roles, new means. Chicago: The National Society for the Study of Education, 1969.

Tyler, R.W. & Wolf, R.M. (Eds.). Crucial issues in testing. Berkeley, California: McCutchan, 1974.

Walsh, W.B. Theories of person-environment interaction: Implications for the college student. Iowa City, Iowa: The American College Testing Program, 1973.

Whimbey, A. Intelligence can be taught. New York: E.P. Dutton & Co., 1975.

Williams, R.I. Black pride, academic relevance, and individual achievement. In R.W. Tyler & R.M. Wolf (Eds.), Crucial issues in testing. Berkeley, California: McCutchan, 1974.

Willingham, W.W. College placement and exemption. Princeton: College Entrance Examination Board, 1974.

Wirtz, W. The boundless resource. A prospectus for an education-work policy. Washington: New Republic Book Company, 1975.